# خوارزميات لرومنة الأسماء العربية

**(الغامدي، منصور (١٤٢٥هـ) خوارزميات لرومنة الأسماء العربية. مجلة جامعة الملك سعود: علوم الحاسب والمعلومات. ١٧: ١-٢٧.)**

**(Alghamdi, Mansour (2005) Algorithms for Romanizing Arabic names. Journal of King Saud University: Computer Sciences and Information. 17: 1-27.)**

**منصور بن محمد الغامدي**

www.mghamdi.com

مدينة الملك عبد العزيز للعلوم والتقنية، الرياض.

١٤٢٥هـ

**ملخص البحث:**

تشكل عملية النقل الكتابي للأسماء العربية من الحرف العربي إلى الحرف الروماني معضلة لكثير من العاملين في مجال رومنة الأسماء العربية كموظفي الجوازات وشركات الخطوط الجوية وإدارات التسجيل في المؤسسات التعليمية والبنوك، وذلك لسببين: السبب الأول: عدم توفر معيار موحد للرومنة، والسبب الثاني: أنه حتى عند توفر هذا المعيار هناك صعوبة في تطبيقه لعدم إلمام هؤلاء الموظفين بتطبيق معيار الرومنة لأسباب بعضها يتعلق بالحرف العربي والآخر بالكتابة بالحرف الروماني. وقد تم تجاوز السبب الأول بعد صدور المعيار الموحد لرومنة الأسماء العربية الذي نتج عن *الندوة العلمية الأولى حول توحيد معايير النقل الكتابي للأسماء العربية: الأبعاد الأمنية*. وتقدم هذه الورقة مجموعة من الخوارزميات الكفيلة بتطبيق المعيار الموحد المشار إليه. وتُعتبر هذه القواعد أساساً لكتابة برمجيات تتولى عملية الرومنة للأسماء العربية آليا.

## ١. المقدمة

لا يجد الكاتب العربي صعوبة تذكر في كتابة أي اسم[١] عربي بالحرف العربي حتى لو كان يسمعه للمرة الأولى. ويعود ذلك لعاملين: الأول، أن الكتابة العربية تقوم على النطق غالباً؛ بمعنى أن الاســم

[١]. يقصد بالاسم في هذه الورقة كامل الاسم، مثل: (١) كمال ، (٢) روضة خريم، (٣) عبد ا ، (٤) محمد نصر الدين. أما الكلمة فهي مفردة واحدة من الاسم، وتكون إما جزءا من الاسم أو كامل الاسم؛ فالاسم في (١) يتكون من كلمة واحدة والاسمين في (٢) و (٣) يتكون كل منهما من كلمتين، والاسم في (٤) يتكون من ثلاث كلمات؛ لاحظ أن الاسم في (٤) هو الاسم الأول للشخص وليس اسمه واسم أبيه.

يكتب كما ينطق وأن العلاقة بين الفونيم[٢] phoneme والقرافيم[٣] grapheme هي علاقة متطابقة إلى حـد بعيد - وهذا عكس لغات أخرى كالإنجليزية، على سبيل المثال. العامل الثـاني، أن الكتابـة العربيـة المعاصرة تقتصر على الحروف[٤] التي غالبا ما تمثل الصوامت[5] consonants دون علامات التشـكيل[٦] التي غالبا ما تمثل الصوائت[٧] vowels مما يعفي الكاتب من التدقيق في الصوائت والتركيز فقط علـى الصوامت. إلا أن الأمر يختلف عند كتابة الاسم العربي بالحرف الروماني وهو مـا يعـرف بالنقـل الكتابي transliteration[٨]؛ إذ نجد تبايناً واضحاً عند كتابة الأسماء العربية بحروف اللغات التي تستخدم الحرف الروماني في الكتابة. وقد أظهرت نتيجة البحث باستخدام محرك البحـث Google أن الاسـم العربي "محمد" - وهو من أكثر الأسماء شيوعا في العالم ومن أقدمها – يظهـر اختلافـا جليـا فـي الحروف المستخدمة لرومنته (الجدول: ١). فرغم قدم استعمال هذا الاسم وكثرة استخدامه إلا أنه يُكتب بطرق متعددة ولكل منها شيوع في الاستخدام مما يبرز التباين في رومنة الأسماء العربيـة وأن هـذا التباين قديم قدم بداية رومنة الاسم العربي نفسه.

***الجدول ١:*** *تكرار ورود النقل الكتابي للاسـم "محمـد" بـالحرف الروماني باستخدام محرك البحث Google.*

| التسلسل | النقل الكتابي | التكرار |
|---|---|---|
| [١] | Muhammad | ٢,٢٨٠,٠٠٠ |
| [٢] | Mohammed | ٢,٠٠٠,٠٠٠ |

٢. الفونيم: هو أصغر وحدة صوتية تغير المعنى في الكلمة ، فالكلمة "الطور" تتكون من ستة فونيمات: /ء ـَـ ط ط ـُـُـ ر/ . ويظهر هنا الفرق بين الكتابة العادية للكلمة السابقة وبين كتابتها صوتيا لتمثيل الفونيمات. في هذه الورقة، عند كتابة رمز أو رموز للدلالة على الأصوات اللغوية – الفونيمات - فإنها توضع بين خطين مائلين //.

٣. القرافيم: أصغر رسم مستقل يستخدم في الكتابة ، فالكلمة السابقة "الطور" تتكون من سبعة قرافيمات: "ا ل ط ّ ُ و ر". وتُكتب الهمزة – كفونيم – في العربية بخمسة أشكال - قرافيمات - مختلفة: "ء ، ئ ، ؤ ، إ ، أ". هذه الرموز الخمسة لها تمثيل صوتي واحد في الدماغ وهو الفونيم /ء/. في هذه الورقة، عند كتابة الحروف للدلالة على الرموز التي تستخدم في الكتابة العربية فإنها توضع بين علامتي اقتباس "".

٤. الحروف العربية عددها ٣٦ حرفا وهي: "ا ، آ ، أ ، إ ، ئ ، ؤ ، ء ، ى ، ة ، ب ، ت ، ث ، ج ، ح ، خ ، د ، ذ ، ر ، ز ، س ، ش ، ص ، ض ، ط ، ظ ، ع ، غ ، ف ، ق ، ك ، ل ، م ، ن ، هـ ، و ، ي". هذه هي جميع الحروف التي نستخدمها في الكتابة.

5. الصوامت: عددها في العربية ٢٨ صامتا، وهي جميع الحروف العربية عدا: الواو المسبوقة بضمة والياء المسبوقة بكسرة. أما الحروف: " ا ، آ، أ ، إ ، ئ ، ؤ ، ى ، ة ". فهي زائدة أي أن هناك رموز غيرها تمثل الصوت نفسه. فعلى سبيل المثال، "ة" إما أن تنطق /ت/ ويمثلها الرمز "ت" أو أن تنطق /هـ/ ويمثلها الرمز "هـ". بمعنى آخر، لا تمثل "ة" صوتا مختلفا عن الصوامت الثمان والعشرين في العربية.

٦. علامات التشكيل: َ ، ِ ، ُ ، ّ ، ْ ، ً ، ٍ ، ٌ .

٧. الصوائت العربية هي: الواو المسبوق بضمة، والياء المسبوق بكسرة، والألف المسبوق بفتحة، والضمة، والكسرة، والفتحة. ويرمز لها صوتيا: / ـُـُـ ، ـِـِـ ، ـَـَـ ، ـُـ ، ـِـ ، ـَـ /، على التوالي.

٨. يُعرف ديفيد كريستال [١] النقل الكتابي transliteration بأنه: "تحويل نظام الكتابة للغة ما إلى نظام آخر للغة أخرى بأن يوضع لكل رمز للكتابة في اللغة الأولى رمز آخر يقابله في نظام الكتابة للغة الهدف." ويعرفه رمزي بعلبكي [٢] بأنه: "استنساخ حروف لغة ما بحروف لغة أخرى".

| | | |
|---|---|---|
| [٣] | Mohamed | ١,٦٠٠,٠٠٠ |
| [٤] | Mohammad | ١,١٥٠,٠٠٠ |
| [٥] | Muhammed | ٣٨٨,٠٠٠ |
| [٦] | Mohamad | ٢٦٤,٠٠٠ |
| [٧] | Muhamed | ٦٩,١٠٠ |
| [٨] | Muhamad | ٤٤,٦٠٠ |

ويعود السبب لهذا التباين في النقل الكتابي للأسماء العربية بالحرف الروماني لعدة عوامل، منها:

١) اختلاف النظام الصوتي sound system العربي عن نظيره في اللغات التي تستخدم الحرف الروماني؛ فعلى سبيل المثال، هناك ٢١ فونيما في العربية ليس لها نظير في الإنجليزية [٣].

٢) لم تكن للمستشرقين والرحالة الغربيين الأوائل الذين كتبوا عن العالم العربي دراية دقيقة باللغة العربية وأصواتها. مما جعلهم يعتمدون في النقل الكتابي للأسماء العربية على نطق المتحدثين باللهجات العامية في تلك العصور ثم اجتهاداتهم الشخصية. لذا نجد كاتب كـ "لورنس" يكتب مدينة "جدة" ٢٥ مرة: Jeddah، و ٦ مرات أخرى: Jidda، ومرة واحدة يكتبها: Jedda في الكتاب نفسه الذي نشره للمرة الأولى عام ١٩٢٦م [٤]، وجميعها تختلف عن الاسم المعتمد في المملكة العربية السعودية وهو Jaddah. هذا الاختلاف في كتابة الأسماء العربية استمر بعد ذلك على مبدأ أنه ليس هناك نظام يمكن وضعه للنقل الكتابي بين الحرف العربي والحرف الروماني؛ فيذكر "لورنس" في الصفحة ٩ من المرجع السابق ما نصه: "لا يمكن نقل الأسماء العربية إلى الإنجليزية نقلا دقيقا، بسبب صوامتها التي تختلف عن صوامتنا [يقصد صوامت اللغة الإنجليزية] وصوائتها التي كصوائتنا تختلف من مكان إلى آخر".

٣) أن اسم العلم العربي قد يأتي مفردا كما في "سليمان" وقد يأتي مركبا كما في "أبو داوود" و "عبد ا " مع وجود مسافة بين الكلمتين المكونة للاسم المركب، وهذا يزيد من صعوبة الرومنة؛ إذ نجد الاسم العربي المركب يرومن مرة بكلمة واحدة وأحيانا بكلمتين منفصلتين وبأشكال مختلفة: Abdu Allah, Abdu-allah, Abdullah.

٤) عند كتابة الأسماء العربية بالحرف العربي فإنه غالبا ما تكتب بالحروف دون علامات التشكيل مما يجعل بعض الأسماء موقع لبس في نطقها كما في "حميد" الذي يمكن أن ينطق "حَمِيد" و "حُمَيد"؛ "سَعَد" و "سَعْد". هذا الاختلاف ينتقل إلى الحرف الروماني فيجعل الاسم الواحد يكتب بأكثر من طريقة.

٥) كثرة معايير النقل الكتابي وتباينها [٥، ٦، ٧]؛ فهناك نظام مكتبة الكونغرس، ونظام الأمم المتحدة، ونظام دائرة المعارف الإسلامية، ونظام قاموس فهر، ونظام المنظمة العالمية

للمقاييس ISO. من أجل ذلك عُقدت *الندوة العلمية الأولى حول توحيد معايير النقل الكتـابي: الأبعاد الأمنية*، بجامعة نايف العربية للعلوم الأمنية عام ١٤٢٤هـ. وأوصـت باسـتخدام نموذج مقترح لكتابة الأسماء العربية بالحرف الروماني [٣].

والنقل الكتابي لأسماء الأعلام لا يقتصر على الأسماء العربية، فهو قائم بين اللغات باختلاف أصولها ونظمها الصوتية وأساليبها في الكتابة. إلا أن النقل الكتابي بين اللغات التي تستخدم الحرف نفسه في الكتابة عادة ما يكون أيسر منه عندما تختلف حروف كتابتها. فالنقل الكتابي أقل تعقيدا بين الفرنسية والألمانية والإنجليزية والفنلندية، على سبيل المثال، لأن جميعها تستخدم الحرف الروماني. وغالبا ما تكون كتابة اسم العلم هي نفسها في أي من هذه اللغات، فنجد اسما مثل: Antonio يكتب كما هو في أي من هذه اللغات سالفة الذكر [٨]. إلا أن الأمر يصبح أكثر تعقيدا عندما يختلف الحرف والنظام الصوتي كما هو قائم بين الإنجليزية والصينية [٩ ، ١٠]، وكما هو كذلك بين العربية والإنجليزية [١١]. أما في اليابانية فرغم اختلاف نظامها الكتابي عن النظام الروماني إلا أن النقل الكتابي للأسماء التي تكتب بالحرف الروماني إلى اليابانية يكون باستخدام ألفبائية خاصة بالنقل الكتابي غير تلك التي تستخدم في كتابة اللغة اليابانية، وهذا يسهل إلى حد ما عملية النقل الكتابي [١٢]. وتختلف نظم كتابة اللغات، فحيث نجد أن الكتابة العربية فونيمية، حيث لكل صوت – فونيم - رمز كتابي – قرافيم. إلا أننا نجد لغات أخرى كالصينية لها نظام مقطعي syllabic أي أن لكل مقطع رمز خاص به. فعلى سبيل المثال، لكل مقطع من المقاطع التالية رمز مستقل خاص به: /كـِـ/ ، /كـَـ/ ، /كـُـ/ ويقاس على ذلك الاحتمالات الأخرى المكونة للمقاطع من صوامت وصوائت مما يجعل عدد القرافيمات يصل إلى ما يقرب من ستة آلاف رمز [١٠] وهذا يزيد من تعقيد النقل الكتابي بين هذين النوعين من اللغات إلا إذا استحدثت ألفبائية لغرض النقل الكتابي كما هو بالنسبة لليابانية.

وقد نهج الباحثون طرائق شتى في عملية النقل الكتابي تختلف بناء على اللغات التي يتعاملون معها والأسلوب الذي يرونه أنسب للخروج بما يرونه حلولا ناجعة لما بين أيديهم من بيانات لأسماء أعلام أو مصطلحات. فقام "جارمو توي فونين" وآخرون باستخدام القوانين التحويلية transformation rules في عملية النقل الكتابي لأسماء الأعلام والمصطلحات بين لغات جميعها يستخدم الحرف الروماني كالألمانية والأسبانية والسويدية والفنلندية والفرنسية إلى الإنجليزية [٨]. حيث تستخلص القوانين التحويلية آليا من القواميس. فعلى سبيل المثال، ia في الأسبانية يقابلها y في الإنجليزية كما في "علم الجسد" somatologia و somatology، على التوالي. فيكون القانون التحويلي هنا هو: ia ← y. ولم تكن نسبة الدقة باستخدام هذه الطريقة في النقل الكتابي عالية. فعلى سبيل المثال، نسبة النقل الكتابي الصحيح من الفنلندية إلى الإنجليزية ٧٢%. ويعود السبب في ذلك إلى كثرة الحالات التي ليس لها قانون ثابت كما في: "وزير المالية" بالفنلندية kansleri وبالإنجليزية chancellor.

واستخدم باحثون آخرون القوانين التي تتولى تحويل القرافيمات في لغة كالصينية إلـى رمـوز صوتية ثم تحويلها بعد ذلك إلى حروف رومانية [١١]. فالقوانين هنا ضرورية حتى يكون النقل الكتابي

أقرب ما يكون لنطق الصوت بلغته الأصلية، وحتى تفك رموز الكتابة الصينية التي تقوم على الكتابة المقطعية ولا تعتمد على الألفبائية كما هي الحال بالنسبة للحرف الروماني.

ومن طرق النقل الكتابي الإحصاء [٩، ١٣] الذي يقوم على عامل الاحتمالية probability. وهذه الطريقة كسابقتها تعتمد على قواعد البيانات، إلا أنها في هذه الحالة، يكون هناك أكثر من شكل للنقل الكتابي للكلمة الواحدة. وتعتمد هذه الطريقة على الشكل الأكثر شيوعا. وتكون قاعدة البيانات إما قواميس إلكترونية أو استخدام محركات البحث على شبكة الإنترنت. فعلى سبيل المثال، الاحتمالية المرشحة للنقل الكتابي للاسم "محمد" المذكور في الجدول: ١، هي Muhammad، فاحتماليتها تساوي P=0.29، وهي أعلى من احتمالية ورود الأشكال الأخرى، حيث احتمالية الكلمة التي تليها Mohammed هي P=0.25. ويعاب على هذه الطريقة أن الكلمة الأكثر شيوعا ليست دائما هي الصحيحة. فعلى سبيل المثال كلمة "مكة" وردت رومنتها Makkah ١٦٠,٠٠٠مرة عند البحث عنها في Google بينما وردت بالشكل Mecca ١,١٥٠,٠٠٠ مرة، رغم أن الرومنة الأولى هي الصحيحة والمعتمدة.

ومن طرق رومنة الأسماء العربية القائمة الآن ما تعتمد على قواعد البيانات أي أنه يتم إدخال الاسم يدويا بالحرفين العربي والروماني للمرة الأولى وعند وروده فيما بعد فإن النظام يستدعي الاسم الذي سبق تخزينه في قاعدة البيانات [١٤]. وفي هذه الحالة فإنه لا بد أن يكون مدخل الاسم على دراية بقواعد رومنة الأسماء العربية وعلى دراية كذلك بنطق الاسم. إلا أن هذا النوع من النظم يعتمد على اجتهاد مدخل البيانات مما يجعل الرومنة عرضة للخطأ إضافة إلى أنها لا تفرق بين الأسماء عند تطابقها في الحروف واختلافها في التشكيل كما في: "حَسِيب" و "حُسَيب" وغيرها من الأسماء المتشابهة [١٥]. وفي كل مرة يرد اسم جديد يتطلب من يرومنه ويضيفه إلى قاعدة البيانات.

عالج "سنها" [١٦] أسماء الأعلام على أنها كلمات مجهولة unknowns وعرفها بأنها أحد أنواع الكلمات التي لا يحويها المعجم؛ بمعنى أن الأسماء من الكلمات التي ليس لها معنى ومن ثم لا بد من معالجتها عند الترجمة بأن تنقل كتابيا ولا تترجم. ففي مرحلة التحليل الصرفي للكلمات أثناء عملية الترجمة الآلية يستبعد المحلل الصرفي الكلمات التي ليس لها جذور في قاعدة بيانات اللغة وتعامل هذه الكلمات على أنها أسماء تنقل كتابيا ولا تترجم.

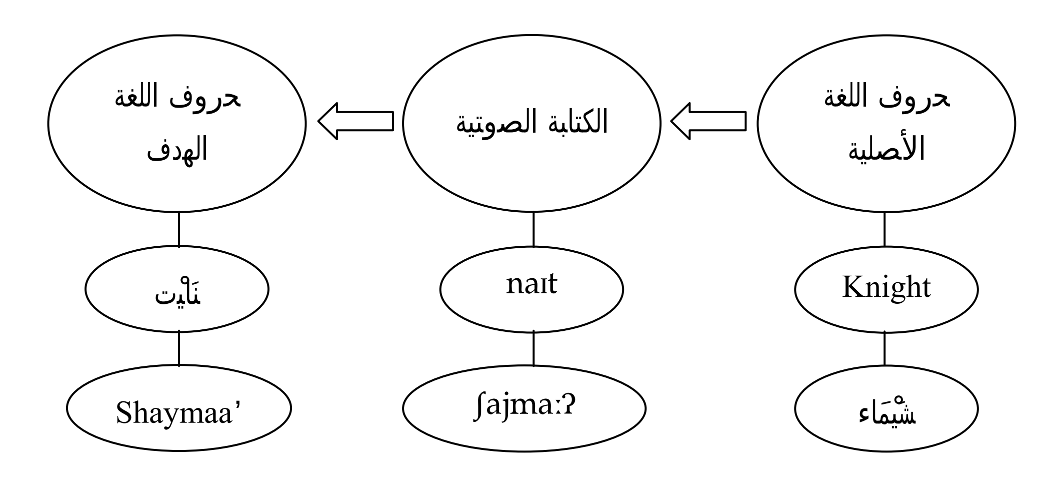
وقد استخدمت برمجيات مختلفة لتنفيذ عملية النقل الكتابي؛ ومنها Hidden Markov Model التي استخدمت في تحويل حروف أسماء الأعلام من الإنجليزية إلى الكورية التي تتطابق فيها الأصوات مع الحروف [١٧]. كما استخدمت الشبكة العصبية الصناعية artificial neural networks في تشكيل الأسماء العربية تمهيدا لرومنتها [١٨].

وبغض النظر عن طريقة النقل الكتابي المتبعة، فإن جميعها لا بد أن تحوي عنصر نطق الاسم بلغته الأصلية وليس فقط كتابته وقد أشارت جميع المراجع السابقة إلى استخدامها لهذا العنصر بشكل

أو بآخر. هذا يعني أن عملية النقل الكتابي تتكون من ثلاثة عناصر أساسية لا بد من توفرها فـي أي نظام للنقل الكتابي (الشكل: ١):

**العنصر الأول:** الاسم مكتوب بحرف اللغة الأصلية، وهو الاسم المراد نقله كتابيا. هذا الاسـم إمـا أن يدخل يدويا أو يكون ضمن قاعدة بيانات لأسماء الأعلام.

**العنصر الثاني:** الكتابة الصوتية، وفيها يتم تحويل الاسم من الكتابة بلغته الأصلية إلى كتابـة صـوتية تقوم على نطق الاسم وليس على كتابته. وهذه العملية تعـرف باسـم "مـن القـرافيم إلـى الفـونيم" grapheme-to-phoneme. فعلى سبيل المثال، الحرف "s" في الاسم الإنجليزي المشهور Charles ينطق /z/ وليس /s/ ومن ثم يكتب صوتيا حسب نطقه. وعندما يكتب بالحرف العربي فإنه يكتب "تشـارلز"، ويلاحظ أن "s" هنا لا يكتب مقابلها "س" عند كتابة هذا الاسم بالحرف العربي ولكن "ز"، وذلك لأنهـا تنطق هكذا بلغة الاسم الأصلية. وعند كتابة الاسم Knight بالعربية فإنه يكتب "نايـت" فالقرافيمـات المستخدمة هنا في الكتابة لا تؤخذ في الحسبان؛ وإنما النطق، لذا تحذف الحروف اللاتينية: k g h لأنها لا تنطق وإلا كان النقل الكتابي لهذا الاسم هكذا: "كناغهت"، وهذا لا يحدث. وغالبا ما تكون الرمـوز المستخدمة في الكتابة الصوتية هي الألفبائيـة الصـوتية الدوليـة International Phonetic Alphabet ويرمز لها اختصارا IPA وتضم أكثر من مئة رمز تمثل جميع أصوات اللغات البشرية [١٩]. ويمكن بهذه الألفبائية كتابة أي اسم بغض النظر عن لغته.




*الشـكل: ١. مراحل النقل الكتابي مع التوضيح بمثال للنقل الكتابي لاسم مـن الإنجليزية إلى العربية، وآخر من العربية إلى الإنجليزية.*

**العنصر الثالث:** كتابة الاسم بحروف اللغة الهدف أو ما يعرف بـ "من الفونيم إلى القرافيم" phoneme-to-grapheme. ويتم في هذه المرحلة تحويل الكتابة الصوتية إلى قرافيمات اللغة الهدف، حيـث يـتم اختيار القرافيم الذي صوته يمثل أقرب نطق لصوت رمز الكتابة الصوتية بحيث يكون هناك أكبر قدر

ممكن من التشابه إما في مخرج الصوت أو في طريقة إخراجه أو فيهما معا. فعلى سبيل المثال، عند النقل الكتابي لاسم المدينة: Reading إلى العربية فإننا نكتبه: "ريدنغ"، ذلك لأن مخرج الصامت الأخير في هذا الاسم /ŋ/ من اللهاة حيث يشترك مع الصامت العربي /غ/ في نفس المخرج ولو أن الأول أنفي والآخر فمّي رخو، إلا أن هذا هو الصوت الأقرب له.

ولا بد من الإشارة إلى أن عملية التحويل من الفونيم إلى القرافيم ليست آلية صرفة بمعزل عن خصائص اللغة الهدف، إذ لا بد أن يخضع النقل الكتابي للقوانين الفونولوجية phonological rules والفونوتكتيكية[٩] phonotactic rules للغة الهدف؛ فالاسم Clinton يحتوي على صامتين في بداية الكلمة – المقطع الأول – ويكتب صوتيا /klıntn/[١٠] وعند نقله كتابيا إلى العربية فإننا نكتبه /كِـلِـنْـتُـون/ وهكذا ننطقه. وقد أدخلنا الصائت /ـِـ/ بين /ك/ ، /ل/ رغم أنه غير موجود في الاسم بلغته الأصلية وذلك لكون اللغة العربية لا تسمح بورود صامتين متتاليين في بداية المقطع؛ وهذا ما يعرف بالساكن في بداية الكلمة. وأدخلنا الصائت الطويل قبل الحرف الأخير لأن اللغة العربية لا تجيز ورود ثلاثة صوامت متتالية في أي موقع من المقطع.

## ٢. الطريقة

تتولى هذه الورقة وضع القواعد التحويلية للأسماء العربية من الحرف العربي إلى الروماني بتطبيق معايير النقل الكتابي المشار إليها في *الندوة العلمية الأولى حول توحيد معايير النقل الكتابي لأسماء الأعلام العربية: الأبعاد الأمنية* (الملحق: ١).

سيكون هناك تجاهل لبعض القوانين الفونولوجية الواردة في [٢٠] المتعلقة بتحويل الكتابة العادية إلى كتابة صوتية، ذلك لكون النقل الكتابي تمثيلا للقرافيمات مع الأخذ في الاعتبار عامل النطق. فعلى سبيل المثال، عند تمثيل الاسم "عَنْبَر" فإن قانون الإقلاب لا يطبق هنا ولا يتحول الاسم إلى "عَمْبَر" ليصبح عند الرومنة 'Ambar'. وإنما يبقى مطابقا للحروف العربية فيكتب 'Anbar'. ويترك عامل تأثير الأصوات على بعضها للغة الهدف. ففي الإنجليزية – على سبيل المثال – يتحول /n/ إلى /m/ عندما يتبعه /b/:

$$n \rightarrow m / - b$$

كما في كلمة fan belt التي تنطق /fam-belt/، هذا يعني أن هذا القانون الفونولجي موحد في اللغتين الإنجليزية والعربية، على الأقل.

---

٩. القوانين الفونولوجية: خاصة بأصوات اللغة phonemes ونظامها الصوتي sound system. والقوانين الفونوتكتيكية: خاصة بتسلسل عناقيد أصوات اللغة؛ فاللغة العربية لا تسمح إلا بورود صامت واحد في بداية المقطع ولكنها تسمح بتسلسل صامتين في نهايته، إضافة إلى أنها لا تسمح بتسلسل بعض الصوامت مثل /ع/ ، /ح/. ونجد في المقابل أن نظام عناقيد الصوامت في اللغة الإنجليزية يسمح بتسلسل يصل إلى حد ثلاثة صوامت في بداية المقطع كما في street.

١٠. لاحظ أن الحرف "o" لا ينطق وإنما يتحول الحرف الأخير "n" في الاسم إلى ما يعرف بالحرف المقطعي syllabic.

وحتى تتم عملية النقل الكتابي بشكل دقيق مع الأخذ في الحسبان جميع الاحتمالات الممكنة بما في ذلك إدخال الكلمات العربية، فقد قُسمت عملية النقل الكتابي إلى أربع مراحل:

**المرحلة الأولى: معالجة النص العربي:**

وتتم في هذه المرحلة عملية مراجعة القرافيمات المدخلة والتأكد من الكتابة السليمة لها خاصة وأن من المستحيل ملاحظة الأخطاء في كتابة بعض القرافيمات العربية. فعلى سبيل المثال، لا يمكن إدراك تكرار طباعة حركة بعينها على الحرف نفسه، فلو تكررت طباعة الفتحة على حرف الكاف في كلمة "كَتب" عشرات المرات فإنها لا تغير من شكل الكلمة في شيء. إلا أنه عند معالجة الكلمة أو محاولة رومنتها فإن البرمجيات ستنظر إلى كل مرة أدخلت فيها الفتحة على أنها شكل مستقل character. ولهذا السبب ولأسباب تتعلق بالخط العربي وكتابته كالمد في كلمة مثل "أحمـــــــد"، فإن هذه المرحلة ضرورية لتهيئة النص للمرحلة التالية وهي الكتابة الصوتية. والقواعد المطلوب تنفيذها في هذه المرحلة هي:

١) وضع قاعدة بيانات بالأسماء المستثناة في كتابتها بالعربية، وتكتب صوتيا؛ مثل:

أ) "ا " ← اللاه

ب) "طه" ← طَاهَا

ت) "يسين" ← يَاسِين

٢) التأكد من أن جميع الحروف مشكلة أي أن كل حرف متبوع بأحد الحركات التالية فقط:

أ) الفتحة: " َ " "أَحْمَد"

ب) الكسرة: " ِ " "سِهَام"

ت) الضمة: " ُ " "بُتُول"

ث) الشدة + الفتحة: " " "سِنمار"

ج) الشدة + الكسرة: " " "سُرَيع"

ح) الشدة + الضمة: " " "سَلوم"

خ) الشدة منفردة " ّ " (فقط في آخر الاسم) "صَفِيّ"

د) السكون: " ْ " "عَدْنَان"

والحروف التالية هي الوحيدة المستثناة من التشكيل، وفي حالة وجودها مشكلة فيحذف تشكيلها:

- "ا" "سَمَاء، السبْت، لِمَا"
- "آ" "آمَال"
- "و" إذا كان مسبوقا بضمة "عَاُوش"
- "ي" إذا كان مسبوقا بكسرة "عَلِيم"

- "ل"       إذا كان مسبوقا بـ "ا" في بداية الكلمة ومتبوعا بحرف عليه شدة. "السمَاء"

٣) يعتبر التشكيل التالي غير صحيح ويعدل:

أ) وجود أكثر من تشكيل على الحرف الواحد سواء كان من نفسه كفتحتين أو مختلفا كضمة يتبعها سكون، ويستثنى من ذلك الشدة وما يتبعها من حركات.

ب) وجود التنوين على حرف غير الحرف الأخير من الكلمة، فالخطأ هنا في إدخال التشكيل؛ إذ لا يكون التنوين إلا في نهاية الكلمة.

ت) إذا سبقت الشدة حركة. فالأصل أن تتبع الحركة الشدة ولا تسبقها.

٤) تحذف مدود الحروف "ـ" ، كما في "محـــــــــــــــمد" التي تصبح "محمد".

٥) تحذف جميع علامات التشكيل – عدا الشدة - على الحرف الأخير من كل كلمة.

٦) يحذف التشكيل لو أتى قبل الحرف الأول من الكلمة.

٧) تضاف ضمة " ُ " على آخر حرف في كل كلمة من الأسماء المركبة عدا الكلمة الأخيرة ("شمسُ الدين").

**المرحلة الثانية: قواعد الكتابة الصوتية "من الفرافيم إلى الفونيم":**

ويتم في هذه المرحلة تحويل الكتابة العادية بالحرف العربي إلى كتابة صوتية باستخدام الحرف العربي أيضا، ذلك لأن الحرف العربي يحتوي على جميع الفونيمات العربية ومن ثم ليس هناك ضرورة لاستخدام الرموز الصوتية الدولية IPA.

١. "ل" ← ∅ / إذا كان غير مشكل ("الزامِل" ← /ءَززَـَمِل/)

٢. " ْ " ← ∅ ("شَهْد" ← /شَهد/)

٣. " ّ " ← يكرر الحرف الذي يسبقها ("عَبّاد" ← / عَببَـَد /).

٤. "أ " ← "ء " ("أحْمَد" ← /ءَحمَد/)

"إ" ← "ء " ("إبْرَاهِيم" ← /ءِبرَـَهِـِم/)

"ؤ" ← "ء" ("لُؤَي" ← /لُءَي/)

"ئ" ← "ء " ("رَائِد" ← /رَـَءِد/)

٥. "آ" ← "ء ـَ ـَ" ("مَآثِر" ← /مَءَـَثِر/)

٦. "ا" ← ∅ / إذا كان في بداية كلمة مسبوقة بكلمة أخرى ("نَصْرُ الدين" ← /نَصرُ ددِـِن/)

← "ـَ" / في أي مكان آخر ("صَالِح" ← /صَـَلِح/)

٧. "ة" ← "هـ" / في آخر الاسم ("فَاطِمَة" ← /فَـَطِمَه/)

← "تُ" / في أي مكان آخر ("رَحْمَةُ ا " ← /رَحمَتُ للـَـَه/)

٨. "و" ← "ـُـ" / إذا كان مسبوقا بضمة ("عُهُود" ← /عُهُد/)
٩. "ي" ← "ـِـ" / إذا كان مسبوقا بكسرة ("خَمِيس" ← /خَمِـِس/)
١٠. "ى" ← ∅ ("شَذَى" ← /شَذَ/)[11]

**المرحلة الثالثة: قواعد الرومنة؛ "من الفونيم إلى القرافيم":**

تطبق في هذه المرحلة القواعد التي خلصت إليه *الندوة العلمية الأولى حول توحيد معايير النقل الكتابي: الأبعاد الأمنية* (الملحق: ١). وتتم في هذه المرحلة تحويل كامل الحروف العربية إلى حروف رومانية.

**المرحلة الرابعة: قوانين ما بعد الرومنة:**

تقوم هذه المرحلة بمعالجة النقل الكتابي للأسماء حتى ينسجم مع كتابة الأسماء بالحرف الروماني، مع تطبيق بعض القواعد الخاصة بعملية النقل الكتابي كحروف الصوائت الطويلة والفصل بين أداة التعريف والاسم. وتتكون من القواعد التالية:

١. تدمج الكلمات المكونة للأسماء المركبة في كلمة واحدة: "عبد ا " تصبح ‘Abdullaah
٢. الشرْطة والحروف الكبيرة:
   أ. توضع الشرطة "-" إذا بدأ الاسم الأخير بعد رومنته بـ "a" متبوعا بـ:
   - الحرف نفسه مكرر، كما في: as-salim ← assaalim.
   - حرفين مكررين، كما في: ash-shamsaan ← ashshamsaan .
   - الحرف الروماني "l"، كما في: al-muslim ← almuslim.
   ب. قواعد تكبير الحروف capitalization:
- يكون الحرف الأول في الكلمة حرف كبير إذا لم يكن في الكلمة شحطة "-" ولا تكن الكلمة aal، كما في: آل مسلم التي ترومن aal Musallam.
- الحرف الأول الذي يلي الشحطة إن وجدت في الاسم، كما في: as-Saalim و ash-Shamsaan.
- الحرف الأول الذي يلي رمزي الهمزة "’" أو العين "‘" إذا أتيا في بداية الكلمة، كما في: "أحمد" ترومن ’Ahmad و "علي" ترومن ‘Alee.

٣. تحويل الحروف التالية:

---

[11]. من توصيات ندوة النقل الكتابي أن تحول الألف المقصورة والفتحة التي تسبقها إلى صائت قصير a حتى يمكن التمييز بينها وبين الألف الممدودة كما في "شذا" التي يصبح فيها الألف والفتحة التي تسبقها aa.

أ. ii ← ee كما في: Hamiid ← Hameed

ب. uu ← oo كما في: Humuud ← Humood

وللتأكد من عمل الخوارزمات السالفة الذكر فقد قام الطالب / منصور الوائلي بكتابة برمجيات حاسوبية لتنفيذها باستخدام لغة الفيجوول بيسك Visual Basic ، ويمكن بواسطة هذه البرمجيات تحويل كتابة الاسم العربي من الحرف العربي إلى الحرف الروماني (الشكل: ٢.). البرمجيات متوفرة على الرابط: http://www.mghamdi.com/A2R.zip ويمكن تنزيلها واستخدامها.

ومما تقوم به البرمجيات المذكورة معالجة الأسماء العربية قبل رومنتها (الشكل: ٣.) ومعالجتها كذلك بعد الرومنة كأن تبدأ الاسم بحرف كبير. ويمكن كذلك الإطلاع على كامل شفرة البرمجيات source code بالدخول على الرابط: http://www.mghamdi.com/A2R_SC.zip.




***الشكل: ٢.*** *واجهة الاستخدام أثناء تحويل الاسم العربي "فيصل" إلى Faysal.*

```
Form1.frm - Notepad
File  Edit  Format  View  Help
Private Sub first_litter_validation()
On Error Resume Next
' "َ" "ِ"كسره "ُ"ضمة "ٌ"ضمتين "ً"فتحتين "ٍ"كسرتين "ّ"شدة '
If j = 1 And (nam(j) = "َ" Or nam(j) = "ِ" Or nam(j) = "ُ" Or nam(j) = "ٌ" Or nam(j) = "ً" Or nam(j) = "ٍ" Or nam(j) = "ّ"
Call remove(j)
b = True
Exit Sub
End If
8 b = False
End Sub
Private Sub diacritical_validation()
On Error Resume Next
'if the litter is vowelized and not ا or آ
If (nam(k) = "َ" Or nam(k) = "ِ" Or nam(k) = "ُ" Or nam(k) = "ٌ" Or nam(k) = "ً" Or nam(k) = "ٍ" Or nam(k) = "ّ" Or na
If j = 1 Then "GG
b = False
Exit Sub
Else
```

***الشكل: ٣.*** *تحتوي شفرة البرمجيات على معالجة للأسماء العربية وتشكيلها قبل رومنتها.*

## ٣. النتائج والتوصيات

عرضت هذه الورقة مجموعة من الخوارزميات الكفيلة بتحول الأسماء العربية من الحرف العربي إلى الحرف الروماني. ويمكن استخدامها كذلك مع بعض التعديلات لوضع خوارزميات معاكسة تقوم بتحويل الأسماء العربية من الحرف الروماني إلى الحرف العربي. ولتطبيق الخوارزميات المعروضة في هذه الورقة ومعرفة المدخل والمخرج لبعضها، فإن الجدول: ٢ يعرض لبعض الأسماء العربية والمراحل التي تمر فيها عند تطبيق الخوارزميات المذكورة لتحويل هذه الأسماء من الحرف العربي إلى الحرف الروماني.

***الجدول ٢:*** *تطبيق قواعد الرومنة على أربعة أسماء عربيـة. يرمـز "م" إلـى المرحلة ويرمز الرقم الذي يليه إلى رقم القاعدة تحت تلك المرحلة.*

| القاعدة | الاسم ١ | الاسم ٢ | الاسم٣ | الاسم ٤ |
|---|---|---|---|---|
| | عَبْدُ القَدِير | حَمـــاد | رُكْن الدين | عَلِي |
| م ١: ٤ | | حَماد | | |
| م ١: ٥ | عَبْد القَدِير | | رُكْن الدين | |

| القاعدة | الاسم ١ | الاسم ٢ | الاسم٣ | الاسم ٤ |
|---|---|---|---|---|
| | عَبْدُ القَدِير | حَمـ ـ ـاد | رُكْن الدين | عَلِي |
| م ١: ٧ | عَبْدُ القَدِير | | رُكْنُ الدين | |
| م٢: ١ | | | رُكْنُ ادين | |
| م٢: ٢ | عَبدُ القَدِير | | رُكنُ ادين | |
| م٢: ٣ | | حَمَّاد | رُكنُ ادِدين | |
| م٢: ٦ | عَبدُ لقَدِير | حَممَد | رُكنُ ددِين | |
| م٢: ٩ | عَبدُ لقَدِ ر | | رُكنُ ددِ ن | عَلِ |
| م٣: | ‘abdu lqadiir | hammaad | ruknu ddiin | ‘alii |
| م٤: ١ | ‘abdulqadiir | | ruknuddiin | |
| م٤: ٢ | ‘Abdulqadiir | Hammaad | Ruknuddiin | ‘Alii |
| م٤: ٣ | ‘Abdulqadeer | | Ruknuddeen | ‘Alee |

وتبين عينة الأسماء في الجدول: ٢ كيف تنتقل الأسماء العربية من مرحلة إلى أخرى إلى أن تصبح في شكلها النهائي بالحرف الروماني، مع ملاحظة أن مخرجات قانون ما هي المدخلة للقانون الذي يليه. بمعنى آخر، أن الخوارزميات المعروضة هنا متسلسلة التطبيق.

ونخلص إلى القول بأننا قد وضعنا في هذه الورقة مجموعة متكاملة من القوانين الكفيلة بتنفيذ عملية النقل الكتابي لأي اسم عربي من الحرف العربي إلى الحرف الروماني. ونقدم في الشكل: ٤ رسما تخطيطيا لنظام حاسوبي مقترح يمكن عند الانتهاء منه أن يقوم بعملية متكاملة لرومنة الأسماء العربية آليا. مما سيسهل على العاملين في مجال رومنة الأسماء العربية - كموظفي الجوازات والبنوك وشركات الخطوط – أداء عملهم بدقة عالية تزيل اللبس والتعددية في رومنة الأسماء العربية في المستندات والبطاقات الدولية كالجوازات وبطاقات الائتمان وتذاكر السفر ورخص قيادة السيارات وغيرها. ويتكون النظام المقترح من العناصر التالية:

١) **معالجة الحرف العربي:** يقارن الاسم المدخل بالأسماء الشاذة في قاعدة البيانات وفي حالة كونه أحدها تعاد كتابته ليكون مطابقا لنطقه، وهو ما أُشير إليه سابقا في الفقرة واحد من المرحلة الأولى. وعند توفر محلل صرفي عربي، فإنه يمكن التأكد من صحة الاسم العربي المدخل.

٢) **مراجعة التشكيل:** التأكد من صحة تشكيل الأسماء وأنه ليس هناك خطأ في تشكيل الاسم المراد رومنته، يؤخذ في الاعتبار ما ورد سابقا تحت المرحلة الأولى: معالجة النص العربي.

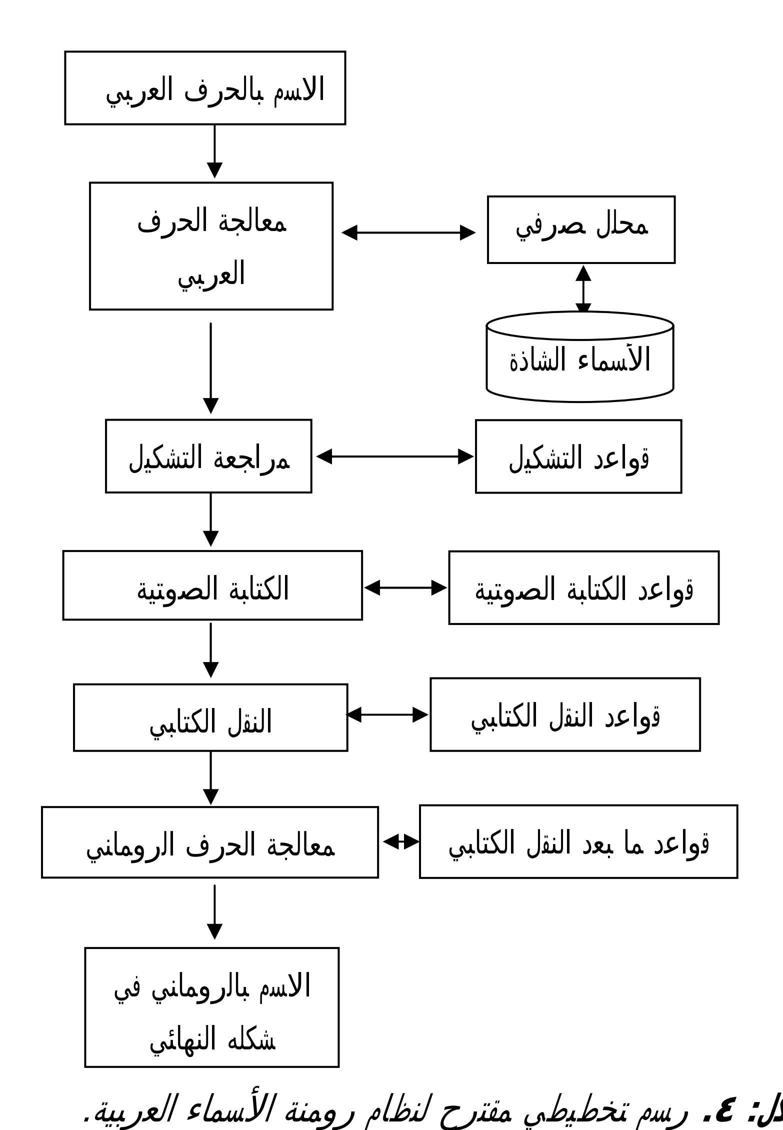
٣) **الكتابة الصوتية:** تحول حروف كتابة الاسم إلى رموز صوتي باستخدام الحرف العربي والقواعد الواردة تحت المرحلة الثانية: الكتابة الصوتية.

٤) **النقل الكتابي:** يحول هنا الحرف العربي إلى ما يقابله من الحـرف الرومـاني حسـب المعيـار المستخدم للنقل الكتابي. وقد استخدم في هذه الورقة ما اتفق عليه المؤتمرون في *النـدوة العلميـة الأول حول توحيد معايير النقل الكتابي لأسماء الأعلام العربية: الأبعاد الأمنية.*

٥) **معالجة الحرف الروماني:** يتم هنا إخراج الاسم بشكله المقبول بالحرف الروماني بما فـي ذلـك تكبير وتصغير الحروف.

صممت هذه العناصر بحيث تكون كل مرحلة مستقلة عن المراحل الأخرى مما يسهل الانتقـال من مرحلة إلى أخرى ويسهل التعديل داخل كل مرحلة حسب المستجدات والتعديلات التي قـد تطـرأ على معايير النقل الكتابي للأسماء العربية في المستقبل.

وختاما فقد قدمت هذه الورقة مجموعة من الخوارزميات تتولى تحويل الأسماء العربيـة مـن الحرف العربي إلى الحرف الروماني آخذة في الاعتبار كل الاحتمالات الممكن ورودها ووضع حـلا لها. وقدمت الورقة كذلك مقترحا لنظام حاسوبي يقوم عند تنفيذه برومنة الأسماء العربية آليا.




***الشكل: ٤.*** *رسم تخطيطي مقترح لنظام رومنة الأسماء العربية.*

**شكر وعرفان:**

يتقدم كاتب هذه الورقة بالشكر والعرفان لوزارة الداخلية ولجامعة نايف العربية للعلوم الأمنية وللمنظمة الدولية للشرطة الجنائية لتنظيمها ورعايتها ندوة توحيد معايير النقل الكتابي. ولمدينة الملك عبد العزيز للعلوم والتقنية الشكر لما قدمته من تسهيلات ودعم لهذه الورقة. وأتقدم بالشكر للدكتور/ عبد الملك السلمان لتفضله بقراءة مسودة هذه الورقة وللملاحظات القيمة التي أبداها مما أنعكس على الشكل النهائي للورقة وما ورد ضمنها.

## المراجع


[١] Crystal, David. A dictionary of linguistics and phonetics. Fifth Edition. Blackwell Publishing. 2003.

[٢] بعلبكي، رمزي. معجم المصطلحات اللغوية. دار العلم للملايين. بيروت. ١٩٩٠م.

[٣] *توصيات الندوة العلمية الأولى حول توحيد معايير النقل الكتابي لأسماء الأعلام العربية: الأبعاد الأمنية*. أكاديمية نايف العربية للعلوم الأمنية. الرياض. ص ٣٨٣. ١٤٢٤هـ.

[٤] Lawrence, T. E. *Seven pillars of wisdom*. Doubleday. 1991.

[٥] القتم، عبد ا . الأنماط المتبعة في كتابة الأسماء الشخصية. *توحيد معايير النقل الكتابي لأسماء الأعلام العربية: الأبعاد الأمنية*. أكاديمية نايف العربية للعلوم الأمنية. الرياض. ٩٧-١١٦. ١٤٢٤هـ.

[٦] القتم، عبد ا . كتابة الأسماء العربية بالحروف اللاتينية. مطبوعات جامعة الكويت. الكويت. ١٩٩٥م.

[٧] صالح، محمود بن إسماعيل. رومنة الأسماء العربية في المراجع الأكاديمية. *توحيد معايير النقل الكتابي بأسماء الأعلام العربية: الأبعاد الأمنية*. ١٧٥-١٩٥. ١٤٢٤هـ.

[٨] Toivonen, Jarmo, Ari Pirkola, Heikki Keskustalo, Kari Visala, and Kalervo Jarvelin. Translating cross-lingual spelling variants using transformation rules. *Information Processing and Management*. in press. 2004.

[٩] Virga, Paola and Sanjeev Khudanpur. Transliteration of Proper Names in Cross-Language Applications. *In Proceedings of SIGIR* 2003, Toronto, Canada.

[١٠] *Meng, H.M.; Wai-Kit Lo; Berlin Chen; Tang, K.*Automatic Speech Recognition and Understanding. ASRU '01. IEEE Workshop on , 9-13 Dec. 2001. 311 – 314. 2001.

[١١] Knight, k. and J. Graehl. Machine transliteration. In Proceedings of ACL. 1997.

[١٢] Stalls, B. and K. Knight. Translating names and technical terms in Arabic text. In *Proceedings of the COLING/ACL Workshop on Computational Approaches to Semitic Languages*. 1998.

[١٣] Abduljaleel, Nasreen. Statistical transliteration for english-arabic cross language information retrieval. In the proceedings of TICIKM, New Orleans, USA. 2003.

[١٤] المحتسب، حسني عبد الغني. نظام آلي لكتابة الأسماء العربية بحروف إنجليزية: نحو معايير موحدة. *توحيد معايير النقل الكتابي بأسماء الأعلام العربية: الأبعاد الأمنية*. ١٧٥-١٩٥. ١٤٢٤هـ.

[١٥] الشمسان، أبو أوس إبراهيم. أسماء الناس في المملكة العربية السعودية. مجلة جامعة الملك سعود: الآداب. ٩: ٣٣١-٤٢٥. ١٤١٧هـ.

[١٦] Sinha, R.M.K. Dealing with unknown in machine translation. IEEE International Conference on Systems, Man and Cybernetics, Volume: 2 :940 - 944 vol.2. 2001.

[١٧] Jung, S. Y., S. Hung and E. Paek. An English to Korean transliteration model of extended Markov window. In proceedings of COLING, Saarbrucken, Germany. 2000.

[١٨] Arbabi, M., S. M. Fischthal, V. C. Cheng and E. Bart. Algorithms for Arabic names transliteration. *IBM Journal of Research and Development*. 38: 183-193. 1994.

[١٩] http://www2.arts.gla.ac.uk/IPA/ipa.html

[٢٠] الغامدي، منصور بن محمد وحسني المحتسب ومصطفى الشافعي. قوانين الفونولجيا العربية. *مجلة جامعة الملك سعود: علوم الحاسب والمعلومات*. ٦: ١-٢٤. ١٤٢٤هـ.

**الملحق: ١**

## نموذج رومنة الأسماء العربية[١١]

١. يكتب الاسم متسلسلا (من الاسم الأول إلى الأب ثم الجد فاللقب) ويكتب رباعيا كلما أمكن ذلك ويوضع ("ابن" أو "بنت" أو "ولد") بين الاسم الأول والثاني وذلك للتفريق بين الأسماء المركبة والمفردة وتكتب عند الرومنة bn, bt, wd.

٢. عدم الأخذ بالإعراب عند كتابة الاسم.

مثال: محمد ← Mu<u>h</u>ammad وليس Mu<u>h</u>ammadu ولا Mu<u>h</u>ammadun

٣. تعامل الأسماء المركبة والكنى معاملة الاسم المفرد ، وذلك بضم مكونات الأسماء المركبة.

مثال: أبو قيس ، شرفُ الدين ← Abuqays, Sharafuddeen

٤. تعامل الـ التعريف حسب نطقها (تختلف الشمسية عن القمرية):

مثال: السالم والمسلم ← as-Saalim, al-Muslim

٥. تعامل آل بطريقة مختلفة عن الـ التعريف:

مثال**:** آل سالم و السالم ← aal Saalim, as-Saalim

٦. تكتب التاء المربوطة حسب نطقها إما هاء أو تاء:

مثال: هبة ، هبة ا ← Hibah, Hibatullah

٧. تكرر الحروف المشددة:

مثال: بَسام ← Bassaam

٨. تعامل الألف المقصورة كالفتحة:

مثال: ليلى ← Layla

٩. تعامل الأسماء بحسب نطقها:

مثال: طه ← <u>T</u>aahaa

١٠. الحروف والحركات:

أ. الحروف العربية التي لها ما يقابلها صوتيا في الحرف الروماني:

| التسلسل | الحرف | | أمثلة بالحرف | |
|---|---|---|---|---|
| | العربي | الروماني | العربي | الروماني |
| ١ | ب | b | بَدْر | Badr |
| ٢ | ت | t | تَمِيم | Tameem |

١١. من توصيات: *الندوة العلمية الأولى حول توحيد معايير النقل الكتابي لأسماء الأعلام العربية: الأبعاد الأمنية*. (١٤٢٤هـ) أكاديمية نايف العربية للعلوم الأمنية. ٣٨١-٣٨٣.

| التسلسل | الحرف | | أمثلة بالحرف | |
|---|---|---|---|---|
| | العربي | الروماني | العربي | الروماني |
| ٣ | ج | j | جَعْفَر | Ja‘far |
| ٤ | د | d | دَاوُد | Daawood |
| ٥ | ر | r | رَامِي | Raamee |
| ٦ | ز | z | زَيْد | Zayd |
| ٧ | س | s | سَامِي | Saamee |
| ٨ | ف | f | فَارِس | Faaris |
| ٩ | ك | k | كَمَال | Kamaal |
| ١٠ | ل | l | لَيْلَى | Layla |
| ١١ | م | m | مُرَاد | Muraad |
| ١٢ | ن | n | نُوح | Nooh |
| ١٣ | هـ | h | هُدَى | Huda |
| ١٤ | و | w | وَلِيد | Waleed |
| ١٥ | ي | y | يَاسِر | Yaasir |

ب. الحروف والحركات العربية التي ليس لها ما يقابلها صوتيا في الحرف الروماني:

| التسلسل | الحرف | | أمثلة بالحرف | |
|---|---|---|---|---|
| | العربي | الروماني | العربي | الروماني |
| ١ | ـَـ | a | سَحَر | Sahar |
| ٢ | ـِـ | i | عِمَاد | ‘Imaad |
| ٣ | ـُـ | u | قُصَي | Qusay |
| ٤ | ـَـا | aa | صَالِح | Saalih |
| ٥ | ـِـي | ee | عِيد | ‘Eed |
| ٦ | ـُـو | oo | مُوسَى | Moosa |
| ٧ | ـَي | ay | شَيمَاء | Shaymaa’ |
| ٨ | ـَو | aw | عَوف | ‘Awf |
| ٩ | ء | ’ | أَحْمد | ’Ahmad |
| ١٠ | ث | th | ثَابِت | Thaabit |
| ١١ | ح | h | حَسان | Hassaan |

| التسلسل | الحرف | | أمثلة بالحرف | |
|---|---|---|---|---|
| | العربي | الروماني | العربي | الروماني |
| ١٢ | خ | kh | خَالِد | Khaalid |
| ١٣ | ذ | dh | ذِيْب | Dheeb |
| ١٤ | ش | sh | شَوقِي | Shawqee |
| ١٥ | ص | s | صَفِية | Safiyyah |
| ١٦ | ض | d | ضُبَيْب | Dubayb |
| ١٧ | ط | t | طَه | Taahaa |
| ١٨ | ظ | z | ظَبْيَة | Zabyah |
| ١٩ | ع | ‘ | عَبْدُ ا | ‘Abdullaah |
| ٢٠ | غ | gh | غَازي | Ghaazee |
| ٢١ | ق | q | قَيْس | Qays |

# Algorithms for Romanizing Arabic Names

**Mansour M. Alghamdi**
*King Abdulaziz City for Science and Technology, Riyadh*

**Abstract.** People such as immigration personnel who work on the Romanization of Arabic names find it troublesome and sometimes confusing. One reason for such difficulty is that an Arabic name can be transliterated in different forms because of the absence of common standards. Another reason is that even if the transliteration of Arabic names is standardized, it is difficult for a layperson to implement it. The first obstacle has been overcome after the *Symposium on Standardizing Arabic Name Transliteration: Security Dimensions* which one of its outcomes was standardized measures for Arabic name transliteration. This paper is to provide algorithms based on the symposium standards that can be used in programming a system to transliterate Arabic names automatically.